

# Minimum reporting standards for plant biology context information in metabolomic studies

Oliver Fiehn · Lloyd W. Sumner · Seung Y. Rhee ·  
Jane Ward · Julie Dickerson · Bernd Markus Lange ·  
Geoff Lane · Ute Roessner · Robert Last ·  
Basil Nikolau

Received: 5 January 2007 / Accepted: 8 June 2007 / Published online: 23 August 2007  
© Springer Science+Business Media, LLC 2007

**Abstract** Plant metabolomics has matured over the past 8 years. Plant biologists routinely use comprehensive analyses of plant metabolites to discover new responses to genetic or environmental perturbation, or to validate initial hypotheses on the function and in vivo action of gene products. The wealth of scientific findings has increasingly provoked interest to share and review raw or processed data from plant metabolomics reports. We here suggest a minimum of parameters to be reported in order to define details of experimental study designs in plant metabolomics studies.

**Keywords** Ontology · Semantics · Minimum information · Metabolite profiling · Metabonomics · Metadata

## 1 Introduction

Four major conferences have been held on the applications and technologies used in plant metabolomics. One of the major observations was that there is no single technology platform that could quantify and identify all plant compounds in a single analysis; instead, technologies are increasingly selected to target specific biological questions which range from evolution of complex traits to adaptation towards abiotic stresses and questions concerning organ- or cell-specific biochemistry. This wealth of scientific findings has increasingly provoked interest to share and review raw or processed data from publications, in order to re-use data to validate preexisting or generate novel hypotheses.

O. Fiehn (✉)  
UC Davis Genome Center, Metabolomics Research Laboratory,  
UC Davis, 451 East Health Sci Drive, Davis, CA 95616-8816,  
USA  
e-mail: ofiehn@ucdavis.edu

L. W. Sumner  
Noble Foundation, Ardmore, OK, USA  
e-mail: lwsumner@noble.org

S. Y. Rhee  
Carnegie Institution, DPB, TAIR, Stanford, CA, USA  
e-mail: rhee@acoma.Stanford.EDU

J. Ward  
Rothamsted Research, Harpenden, Hertfordshire, England  
e-mail: jane.ward@bbsrc.ac.uk

J. Dickerson · B. Nikolau  
Iowa State University, Ames, IA, USA

J. Dickerson  
e-mail: julied@iastate.edu

B. Nikolau  
e-mail: dimmas@iastate.edu

B. M. Lange  
Washington State University, Pullman, WA, USA  
e-mail: lange-m@wsu.edu

G. Lane  
AgResearch, Hamilton, New Zealand  
e-mail: geoff.lane@agresearch.co.nz

U. Roessner  
University of Melbourne, Melbourne, VIC, Australia  
e-mail: ute.roessner@acpfg.com.au

R. Last  
Michigan State University, East Lansing, MI, USA  
e-mail: lastr@msu.edu

What is the baseline of metabolite levels found in model plants such as *Arabidopsis thaliana* or *Oryza sativa*? It has been demonstrated that the metabolomes of higher plants are highly diverse and flexible, spanning an enormous range of complexity and metabolite concentrations. Therefore, answering seemingly simple questions such as ‘how large is the *Arabidopsis* metabolome?’ is still a difficult objective because the answer will depend on the physiological and genotypic settings. For example, which organs were analyzed, what were the illumination and soil conditions used for growing plants, how was the nutrient and watering regime set up, and which genes (if transgenic plants are studied) were altered?

For any single study, these data concerning the generation of the biological materials (i.e. the biological context metadata) are not always presented in detail, according to the requirements for method sections of various plant journals. Initiated by the Metabolomics Society, a number of plant researchers have therefore collaborated to refine the minimum set of required reporting parameters that are essential to describe an experiment. It is important to understand that this effort does not attempt to prescribe experimental designs or distinguish scientifically adequate from inadequate designs. Instead, the proposed standards aim to promote “good” plant biology practices with special provisions to enable comparisons of experimental data and designs electronically and between publications in peer-reviewed journals. In a second but separate step, current best practice standards may be developed, which will go beyond the minimum set of core metadata to be reported and will potentially better reflect the ever changing view of the complement of factors that need to be considered for understanding the metabolome of plants.

The reporting standards proposed here have been reviewed and improved by an in-depth discussion of the participants of the 4th International Conference on Plant Metabolomics, held in Reading, UK, in April 2006. However, despite our best efforts, we may have overlooked important criteria or parameters. In addition, the notion of ‘minimum reporting standards’ cannot refer to an impartial concept but is the result of prolonged discussions to reach consensus. The notion is that ‘minimum reporting standards’ will be endorsed and supported by the plant biology community at large, in order to legitimate mandatory reporting requirements adopted by funding agencies, foundations, scientific organizations and journals. In this respect, the standards presented here do not represent an end point but rather an initial milestone for ongoing discussions. The authors therefore appreciate feedback and constructive criticism which would be incorporated into refined versions of the ‘reporting standards’ documents, that will be available from the Metabolomics Standards

Initiative (MSI) website (<http://msi-workgroups.sourceforge.net/bio-metadata/>). Comments can be also sent to an open list ([Msi-workgroups-feedback@lists.sourceforge.net](mailto:Msi-workgroups-feedback@lists.sourceforge.net)) without the need to subscribe to one of the specific MSI workgroups mailing lists.

## 2 Materials and methods

### 2.1 The Standards generation process

In 2005, the Metabolomics Standards Initiative was formed as result of a workshop organized by the U.S. National Institutes of Health. This Initiative was supported and endorsed by the Metabolomics Society. According to the general metabolomics workflow, one of the key parameters of standardization for reporting metabolomics data was identified as biological context information. It was recognized that study designs and emphases of different fields of biology call for distinct (but small) working groups whose tasks were to compile initial lists of required standards which later should be refined by the larger biology context communities. The active participation of governmental agencies and industrial corporations was actively sought; however, most collaborators were affiliated with public research organizations. The biology subgroups comprised plant biology, mammalian and clinical biology, microbiology, and environmental biology. Group chairs held contact via exchange of documents and teleconferences, organized and chaired by Don Robertson (Pfizer Global Research & Development, Ann Arbor, MI, USA). The chairs initially outlined work plans and exchanged information with other MSI working groups, namely those working on issues of Chemical Analysis, Data Processing & Statistics, Data Exchange, and Ontologies. These exchanges occurred via workshops, conference reports, teleconferences and the MSI website (see also reports of these working groups in this same issue of Metabolomics). The metabolomics society further formed an oversight committee to coordinate activities, led by Oliver Fiehn (UC Davis, USA).

The plant biology context work was founded on previous publications that laid the groundwork for reporting standards. Specifically, the architecture for metabolomics (ArMet) (Jenkins et al. 2004, 2005) and the ‘Standard Metabolic Reporting Structure’ document (SMRS) (Lindon et al. 2005) were utilized as starting points. These evolving standards were complemented by demands for ‘Minimum Information About a Metabolomics Experiment’ [MIA-Met] (Bino et al. 2004); which recognized efforts by other communities. Especially for genomics studies, the need for standardized reporting had been recognized, and several

initiatives have evolved, such as the Minimum Information About a Microarray Experiment [MIAME] (Brazma et al. 2001); the Reporting Structure for Biological Investigations [RSBI] (<http://www.mged.org/workgroups/rsbi/rsbi.html>), the Functional Genomics Ontology [FuGO] (<http://obi.sourceforge.net/>), ‘Chemical effects in biological systems—data dictionary’ [CEBS-DD] (Fostel et al. 2005); and the Proteomics Standards Initiative [PSI] of the Human Proteome Organization.

An initial draft of the document presented here was circulated between the members of the working group, chaired by Basil Nikolau (Iowa State University, Ames, IA, USA). This document was then released and discussed at a 90-min workshop with about 150 participants at the 4th Plant Metabolomics Conference, Reading, UK, in April 2006. Ultimately, the refined and updated version presented here was released for discussion at the 2nd Annual Conference of the Metabolomics Society (Boston, 24–29 June 2006).

### 3 Results

The reporting standards for detailing plant metabolomics studies build upon the commonly accepted practice of reporting plant biological, and specifically, plant physiological experiments. However, the best practice for such reporting has never been formally laid out and enforced by journals. For example, author guidelines in the *Plant Journal* detail that ‘Experimental procedures should be sufficiently detailed to enable the experiments to be reproduced’. The level of experimental detail presented in *Plant Journal* is therefore only subject to the peer-review process, which often focuses on the justification and relevance of the scientific content, rather than methodological aspects. The journal *Plant Cell* is more specific, detailing the author guidelines for ‘Method’ descriptions by the subheadings ‘Large scale experiments’ and ‘Quantification of molecules’. Nevertheless, the instructions in many journals are necessarily more general than specific comments and metadata that are needed for electronic repositories and wide scale re-use of quantitative and qualitative data are not necessarily reported. We have therefore included such classic descriptors of good practice of plant biological experiments and have consequently structured our considerations for a minimum list of core metadata to the following four major classes:

- (i) The description of the physical object under investigation (the ‘biosource’), which includes genotypic and spatial information
- (ii) metadata relating to the (average) growth history of plants, excluding treatments

- (iii) specifications of the physiological or biochemical intervention(s) to which plants were subjected as treatment
- (iv) details of the harvest and post-harvest conditions, in order to assess the conditions at harvest, and likely alterations of metabolic contents due to quenching and post-harvest storage conditions.

Consideration was also given to how the requirement of such minimal metadata should be enforced when metabolomics data are being submitted. For example, although parameters are left sufficiently vague enough to suit many different studies, certain metadata might still be omitted from submission to journals or databases. In such cases, it would be required to state and justify where these omissions in the data occur. Valid reasons for not reporting plant context metadata might be inaccessibility of some data (e.g., in studies relating to plant products for end consumers, or for certain field trials), or intellectual property or commercial restrictions. However, even if exact details cannot be given, authors of plant metabolomics data would be required to give general descriptions. Eventually, data or conclusions must be rejected if omissions of plant biology metadata are so severe that the scientific conclusions cannot be reproduced or understood.

#### 3.1 BioSource

This term refers to the physical objects that were subjected to metabolomic analyses, consisting of a description of the species and genotype and the organ that was sampled, and the bulk quantity of sample. In certain cases, more detailed information may be available such as organ specifications, cell types or subcellular compartments. These metadata are only required and meaningful if sampling methods that allowed such annotations were used. The methods used for sampling should be named in such cases to enable independent evaluation of data.

Details and explanations for required BioSource metadata are given below:

---

**Species** Names of species should be described according to the NCBI taxonomy database (Wheeler et al. 2000; Benson et al. 2000) (<http://pubmedexpress.nih.gov/Taxonomy/taxonomyhome.html/index.cgi>). Plant species need to be named in full and not abbreviated, e.g. *Arabidopsis thaliana*.

All necessary information on taxonomic relationships can be derived from the correct species name and thus does not need to be reported further.

---

continued

Genotype	<p>Subspecies information such as ecotype, cultivar and accession should be described according to authoritative databases such as TAIR (<a href="http://www.arabidopsis.org/">http://www.arabidopsis.org/</a>). In the case of crosses or breeding results, available pedigree information must be given. In the case of transgenic or mutant organisms, name of the gene(s) that are up- or down-regulated should be reported, and the GenBank Accession number(s) for the sequence(s) of the corresponding construct(s), in addition to the parental subspecies background information, should be given.</p> <p>According to standard practice in agronomic genotype nomenclature, genotype description should comprise the author who first described or collected the cultivar, e.g. <i>Medicago truncatula</i> (Gaertn) cv. Jemalong. If available, registrations numbers for agronomic plants should be referenced, e.g. USDA GRIN. The number of backcrosses used in breeding needs to be detailed.</p> <p>In case of plant–pathogen interaction studies or other studies where information on multiple genomes is relevant, such metadata should be given.</p>
Organ	Names of organs and plant structure should be described according to the authoritative database (Katica et al. 2007) maintained by the Plant Ontology Consortium to be found at <a href="http://www.plantontology.org/">http://www.plantontology.org/</a> . All necessary information on organ relationships can be derived from the correct organ name and thus does not need to be reported further.
Organ specification	This should be provided <i>only if</i> such information cannot be detailed by <a href="http://www.plantontology.org/">http://www.plantontology.org/</a> (e.g. description of a part of an organ, the specific location of the organ or a specific tissue of an organ).
Cell type	<p>This should be provided <i>only if</i> such information can be detailed in a meaningful manner, e.g. by cell sorting or dissection. Naming according to the authoritative database maintained by the Plant Ontology Consortium is to be found at <a href="http://www.plantontology.org/">http://www.plantontology.org/</a> under plant_structure ontology.</p> <p><i>Only if</i> such information cannot be located at this source the Cell Ontology maintained at Open Biomedical Ontologies group should be taken, which is to be found at <a href="http://lists.sourceforge.net/lists/listinfo/obo-cell-type">http://lists.sourceforge.net/lists/listinfo/obo-cell-type</a>.</p>
Subcellular location	This should be described <i>only if</i> such information can be detailed in a meaningful manner, e.g. by subcellular fractionation. Naming according to the authoritative database (Gene Ontology Cellular Component) maintained by the Gene Ontology Consortium to be found at <a href="http://www.geneontology.org/">http://www.geneontology.org/</a>
BioSource amount	This refers to the mass (mg fresh weight or mg dry weight), number of cells or other measurable bulk quantities (e.g. protein content).

### 3.2 Growth environment

Many parameters of growth history are identical to all the plants in a given study. Researchers tend to refer to their specific growth environments as ‘standard growth conditions’ because they may not have altered these for a long time, or always use the same growth chambers and illumination conditions. However, environmental parameters are known to be very different between laboratories, and severely affect metabolite levels. Apart from obvious parameters such as (abiotic) stress conditions, even subtle alterations, such as emission spectra of light bulbs used for illumination, may cause differences in overall growth and metabolism. On the other hand, it is known that pulses of fertilizations will be reflected in changes in metabolism, and hence we suggest reporting both amount and timing of the nutritional regime. As guideline, parameters should be reported that can easily be monitored by plant researchers such as the type of growth media and light regimes; however, we intentionally do not suggest distinguishing between set points of growth conditions (such as temperature) and actually achieved parameters (which may have deviated from such set points). It is good laboratory practice to report deviations and fluctuations from controlled growth conditions; however, researchers may not be aware or may not have the instrumentation to monitor these parameters. This is an example of how ‘minimal requirements’ may be distinguished from ‘current best practice’ documents.

This section specifically excludes variation of growth conditions that were part of the experimental design, i.e. factors that were altered with the intention to cause metabolic differences. Such differences in study parameters should be reported as ‘treatment’. Although it cannot be made mandatory, documentation of additional metadata should be regarded as part of best plant biology practice, such as application of agrochemicals or biotic plant protection. When investigating other documents relating to the specifics of plant growth reporting requirements, a document was retrieved that was published by the International Committee for Controlled Environment Guidelines (<http://ncr101.montana.edu/>) in March 2004, detailing the ‘Minimum Guidelines for Measuring and Reporting Environmental Parameters for Experiments on Plants in Growth Rooms and Chambers’ ([http://ncr101.montana.edu/min\\_guidelines.pdf](http://ncr101.montana.edu/min_guidelines.pdf)). While we appreciated the efforts of this committee, we felt that many of the recommendations were rather demanding to be put into practice in current laboratory settings, especially in public research institutions. The intention of documents detailing

minimal reporting standards, including the paper presented here, is to detail enough information to re-use data and to understand the concepts and layout of experimental designs. If minimal requirements ask for a level of detail that is usually not reported by researchers, these can hardly serve as consensus, which would be endorsed and followed by the majority of active investigators. Instead, we suggest such guidelines to be part of ‘best practice’ documents.

Details and explanations for the section ‘Plant Growth’ comprise the following factors:

Growth support	Soil (type, supplier), Agar (type, supplier), Vermiculite (type, supplier), hydroponic system (type, supplier, nutrient concentrations) or other support including cell culture (media, volume, cell number per volume).
Growth location	Field trial (location), climate chamber (size m <sup>3</sup> ), greenhouse (details on accuracy of control of light, humidity and temperature conditions), other location (details on size m <sup>3</sup> , accuracy of control of light, humidity and temperature conditions).
Growth plot design	The way to randomize the different genotype $\times$ environment interactions. Either descriptive or using established nomenclature e.g. latin square.
Light	Light quality, light source model/type, light intensity (best reported as empirically measured at plant height), luminescence (daylight) period (h).  For field trials: average light parameters in growing season. Information on time and location of the field trial enables tracking of more precise information if necessary.
Humidity	Humidity (%) at day and at night.  For field trials: average humidity parameters in growing season. Information on time and location of the field trial enables tracking of more precise information if necessary.
Temperature	Temperature (°C) at day and at night.  For field trials: average temperature (°C) at day and at night in growing season. Information about time and location of the field trial enables tracking of more precise information if necessary.
Watering regime	Amount and time of watering per day.  For field trials: average rain fall in growing season. Information on time and location of the field trial enables tracking of more precise information if necessary.  For hydroponic systems: frequency of solution change.
Nutritional regime	Amount and time of additional nutrients given to plants.
Date(s) of plant establishment	Depending on plant study, such dates could comprise: sowing, germination, transplanting, cutting, grafting or other appropriate time stamps.

continued

Other specific metadata	Plant development stage description should accompany time stamps using established nomenclature (Boyce et al. 2001; Pujar et al. 2006; Palmquist et al. 2006).
	Only if applicable.
	Examples comprise translocation of plants from one chamber to another, or the rotational schema of trays within a climate chamber.  Examples comprise agrochemical or preventive maintenance information that is not part of ‘treatment’ factors.

### 3.3 Treatment

Plant biology study designs can be broadly classified according to *Genotype  $\times$  Environment* interactions, or, for the sake of clarity, alterations of parameters that are denoted here as *BioSource  $\times$  Treatment*. Publicly available and authoritative ‘treatment’ databases that label and detail the variety of treatment factors and their relative hierarchy and dependencies are not yet available in repositories like TAIR or PlantOntology. Hence, without further work on ontologies for such terms, it can only be recommended that terminology is used that is frequently found in plant research journals. In addition, we recommend a broad classification of treatment types (biotic, abiotic and intervention), which need to be complemented by information as to the dose or intensity levels, and time intervals or durations in which treatments were given. However, specific treatments (such as use of elicitors like methyljasmonate, abscisic acid or salicylic acid) is often termed as biotic stress treatment, and hence, there is yet some degree of ambiguity in nomenclature.

Treatment factors	Biotic treatment	E.g. infection (species), herbivore attack (species), competition with other plants (species) <i>or other factors</i>
	Abiotic treatment	E.g. light intensity variations, cold acclimation (temperature), heat stress, drought (description of residual growth support moisture, or quantitative description of reduction in watering regime), water stress, saline stress <i>or other factors</i>
	Intervention treatment	E.g. application of agrochemicals, enzyme inhibitors, hormones, elicitors <i>or other factors</i>
Treatment dose or intensity levels		Depending on treatment factors
Treatment time, time intervals and duration before harvest		Depending on treatment factors and treatment time



### 3.4 Harvest

The harvest determines the set point for stopping metabolism, analogous to sampling time points in related documents on biology context metadata. However, apart from a simple time stamp, further metadata are required. For example, if harvests are reported from different research groups that were using identical plants and growth conditions, and sampling at the same hour of the day, results can still be different: one laboratory may have used a 16-h light periods beginning the illumination period at 06:00 a.m., whereas the other laboratory may have begun illumination at 08:00 a.m. Hence, plants in the first study would be 2 h ahead in their daily period of photosynthesis and starch accumulation, which is known to cause metabolic alterations. Therefore, the time point and duration of harvest should be given relative to the photoperiod.

Another parameter is the age of the plants under study, e.g. the time between seed germination and date of harvest, may not necessarily convey similarity or comparability of growth and thus metabolic status. Even in controlled environments, some plants may flower earlier or grow faster than others, which refer to important turning points in the life cycle of a plant. Therefore, plant growth stages need to be defined relative to standardized growth stages. For some model species, like *Arabidopsis*, such growth stages have been defined in the literature; for other species, nomenclature should be used according to established terminology in plant journals. Most recently, an ontology for standard growth stages has been developed for angiosperms (Pujar et al. 2006), and it is recommended to exploit this resource for detailing plant metabolomic experiments.

Analogous to other ‘biology context’ documents, the method and time at which metabolism was stopped are also important to denote. Metabolites differ vastly in their turnover rates, and some (such as glutathione or NADH) are extremely sensitive to oxidation. Therefore, details on the harvest methods need to be provided to enable the assessment and validation of metabolomic data acquisitions.

Harvest date, time	Harvest time relative to the luminescence cycle. Duration of harvest if relevant to the plant study (e.g. for volatile analysis).
Plant growth stage	It is advised to refer to established literature, e.g. for <i>Arabidopsis</i> (Boyes et al. 2001) and <i>Medicago truncatula</i> (Palmquist et al. 2006); and for general growth stage ontology (Pujar et al. 2006).
Metabolism quenching method	Time after harvest before stopping cellular metabolism. (may be greater than weeks for certain post-harvest physiology experiments, may be less than seconds for assessing high turnover metabolites). Method to stop cellular metabolism

continued

Harvest method	Details of operation to gather the plant organ (sample) Details of pooling of plant tissues for analysis
Sample storage	Operations to store sample (e.g. freeze-drying, grinding) prior to preparation for metabolomic analysis. Duration and temperature of storage before extraction for analysis.

## 4 Discussion

This document presents a first attempt to collate the minimum required metadata for reporting plant metabolomic data. Many of the factors are regarded as classic descriptors for any study in plant biology. However, such metadata are often not explicitly detailed but rather used in reference to previous studies, or as in typical plant research reports, the information may be given scattered in text strings throughout the document. The aim here is to provide a guide for gross descriptors of plant studies, in order to allow comparisons of study designs, and to categorize experiments that will be reported in the literature or in databases. Parameters given here are still rather vague and allow for a number of exceptions and deviations. Specifically, controlled vocabularies, ontologies, and exact definitions for value units and string text restrictions need to be developed in order to implement such minimum requirements into executable, queryable and MSI-compliant databases. For this reason, no UML schema can be given at this point in time. To this end, only for some factors such as ‘species’, ‘plant organs’ ‘growth stages’ and ‘cellular compartments’, clearly defined hierarchies and repositories can be used. It will be important to continue and extend efforts on terminology requirements that are compiled by the MSI Ontology working group (Sansone et al. this issue), and to collaborate further with groups within the Open Biomedical Ontology consortium (OBO, <http://obo.sourceforge.net>). As one of the short-term goals, the list presented here will be incorporated into a ‘minimum information’ checklist (MIBBI, <http://mibbi.sourceforge.net/>), which is envisioned to become a general repository for reporting standards in functional genomics and could be a resource for journal editorial guidelines.

It needs to be emphasized that the list presented here does not imply a sufficient or exhaustive description of plant studies in general, or even for plant metabolomics. Many parameters may require much greater detail if a full reproduction of an experiment is desired. For example, air circulation has not been listed as a parameter, despite its great role in evaporation and transpiration, and hence water availability to plants. In addition, air flow can also be a mechanical stress factor. In this discussion, we have opted

against exhaustive lists of parameters because journal reports are also accepted without such level of detail. Nevertheless, plant metabolomics researchers are highly encouraged to collect additional details about their specific studies, which will eventually translate into a better understanding of plant metabolism by exchange and re-use across laboratories and studies.

**Acknowledgments** Work on this document was partially supported by a research grant MCB-0520140 from the U.S. National Science Foundation.

## References

- Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., Rapp, B. A., & Wheeler, D. L. (2000). GenBank. *Nucleic Acids Research*, 28, 15–18.
- Bino, R. J., Hall, R. D., Fiehn, O., Kopka, J., Saito, K., Draper, J., Nikolau, B. J., Mendes, P., Roessner-Tunali, U., Beale, M. H., Trethewey, R. N., Lange, B. M., Wurtele, E. S., & Sumner, L. W. (2004). Potential of metabolomics as a functional genomics tool. *Trends in Plant Science*, 9, 418–425.
- Boyes, D. C., Zayed, A. M., Ascenzi, R., McCaskill, A. J., Hoffman, N. E., Davis, K. R., & Grolach, J. (2001). Growth stage-based phenotypic analysis of Arabidopsis: A model for high throughput functional genomics in plants. *Plant Cell*, 13, 1499–1510.
- Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., Ansorge, W., Ball, C. A., Causton, H. C., Gaasterland, T., Glenisson, P., Holstege, F. C. P., Kim, I. F., Markowitz, V., Matrese, J. C., Parkinson, H., Robinson, A., Sarans, U., Schylze-Kremer, S., Stewart, J., Taylor, R., Violo, J., & Vingron, M. (2001). Minimum information about a microarray experiment (MIAME) – towards standards for microarray data. *Nature Genetics*, 29, 365–371.
- Fostel, J., Choi, D., Zwickl, C., Morrison, N., Rashid, A., Hasan, A., Bao, W., Richard, A., Tong, W., Bushel, P. R., Brown, R., Bruno, M., Cunningham, M. L., Dix, D., Eastin, W., Frade, C., Garcia, A., Heinloth, A., Irwin, R., Madenspacher, J., Merrick, B. A., Papoian, T., Paules, R., Rocca-Serra, P., Sansone, A. S., Stevens, J., Tomer, K., Yang, C., & Waters, M. (2005). Chemical effects in biological systems—data dictionary (CEBS-DD): a compendium of terms for the capture and integration of biological study design description, conventional phenotypes, and 'omics data. *Toxicological Sciences*, 88, 585–601.
- Jenkins, H., Hardy, N., Beckmann, M., Draper, J., Smith, A. R., Taylor, J., Fiehn, O., Goodacre, R., Bino, R. J., Hall, R., Kopka, J., Lane, G. A., Lange, B. M., Liu, J. R., Mendes, P., Nikolau, B. J., Oliver, S. G., Paton, N. W., Rhee, S., Roessner-Tunali, U., Saito, K., Smedsgaard, J., Sumner, L. W., Wang, T., Walsh, S., Wurtele, E. S., & Kell, D. B. (2004). A proposed framework for the description of plant metabolomics experiments and their results. *Nature Biotechnology*, 22, 1601–1606.
- Jenkins, H., Johnson, H., Kular, B., Wang, T., & Hardy, N. (2005). Toward supportive data collection tools for plant metabolomics. *Plant Physiology*, 138, 67–77.
- Ilic, K., Kelloogg, E. A., Jaiswal, P., Zapata, F., Stevens, P. F., Vincent, L. P., Avraham, S., Reiser, L., Pujar, A., Sachs, M. M., Whitman, N. T., McCouch, S. R., Schaeffer, M. L., Ware, D. H., Stein, L. D., & Rhee, S. Y. (2007). The plant structure ontology: a unified vocabulary of anatomy and morphology of a flowering plant. *Plant Physiology*, 143, 1–13.
- Lindon, J., Nicholson, J., Holmes, E., Keun, H., Craig, A., Pearce, J., Bruce, S., Hardy, N., Sansone, S., Antti, H., Jonsson, P., Daykin, C., Navarange, M., Beger, R., Verheij, E., Amberg, A., Baunsgaard, D., Cantor, G., Lehman-McKeeman, L., Earll, M., Wold, S., Johansson, E., Haselden, J., Kramer, K., Thomas, C., Lindberg, J., Schuppe-Koistinen, I., Wilson, I., Reily, M., Robertson, D., Senn, H., Krotzky, A., Kochhar, S., Powell, J., van der Ouderaa, F., Plumb, R., Schaefer, H., & Spraul, M. (2005). Summary recommendations for standardization and reporting of metabolic analyses. *Nature Biotechnology*, 23, 833–838.
- Palmquist, D., & Vance, C. P. (2006). A standardized method for analysis of *Medicago truncatula* phenotypic development. *Plant Physiology*, 142, 207–219.
- Pujar, A., Jaiswal, P., Kellogg, E. A., Ilic, K., Vincent, L., Avraham, S., Stevens, P., Zapata, Z., Reiser, L., Rhee, S. Y., et al. (2006). Whole plant growth stage ontology for angiosperms and its application in plant biology. *Plant Physiology*, 142, 414–428.
- Wheeler, D. L., Chappey, C., Lash, A. E., Leipe, D. D., Madden, T. L., Schuler, G. D., Tatusova, T. A., & Rapp, B. A. (2000). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, 28, 10–14.